

The Depth of Distractor Processing in Search with Clutter

MARY J. BRAVO¹, HANY FARID²

Some search tasks involve looking for a category target in clutter. This is the task faced, for example, by a baggage screener looking for weapons in a suitcase. Such tasks presumably involve the segmentation and recognition of the target object, but it is unknown whether they also involve the segmentation and recognition of the distractor objects. To examine the depth of distractor processing in this task, we had observers search through cluttered displays composed of normal and chimerical distractors. The normal distractors were photographs of recognizable objects, while the chimerical distractors were created by interchanging parts between the normal objects. The observer's task was to identify the display quadrant that contained an animal or a vehicle target. We varied the difficulty of the search task by varying target and distractor discriminability, target uncertainty and target occlusion. Only when the targets were partially occluded did we find an effect of distractor type. In this case, observers may have found the target through a process of mentally eliminating whole distractor objects. When the target was unoccluded, we found no evidence that observers selected and rejected whole distractors during search. This second result supports our previous claim that often the items for search in clutter are not whole objects.

Visual search Clutter Object recognition

¹Department of Psychology, Rutgers University, Camden NJ 08102. Email: mbravo@camden.rutgers.edu; Tel: 856.225.6431; Fax: 856.225.6602

²Department of Computer Science and Center for Cognitive Neuroscience, Dartmouth College, Hanover NH 03755.

1 Introduction

Some search tasks involve finding a target that is specified only by its membership in a broad category, such as "food", "vehicle" or "weapon". Background clutter can have a profound effect on this type of search. This effect arises because clutter interferes with object segmentation, and this interference has two consequences for category search. First, clutter affects the process of target recognition. When objects are presegmented, recognition is likely based on shape (Ullman, 1997). By preventing the segmentation of whole objects, clutter obscures an object's shape and this forces the observer to adopt a different strategy for target recognition (Bravo & Farid, 2006). Second, clutter can affect the process of distractor rejection. When distractors are pre-segmented, as they usually are in a visual search experiment, the distractor objects are likely processed as a single item. But when these same search items are jumbled together, as they are in a cluttered drawer or suitcase, the effective number of search items may change. For example, a distractor object with a salient internal boundary may be segmented along this boundary and processed as two separate items.

The evidence for this second claim, that clutter can change the way distractors are parsed and processed, comes from a recent experiment (Bravo & Farid, 2004). Because the current study builds directly on this earlier experiment, we describe it in some detail here. In the earlier experiment, observers searched for food items among two types of distractor objects. One type, the "simple distractors", had uniform color and texture and weak internal boundaries (e.g., a red fire hydrant). The other type, the "compound distractors", had parts with different colors or textures and strong internal boundaries (e.g., a table lamp with a white paper shade and a blue ceramic base). We assumed that these two types of distractor objects would differ in their segmentability in clutter. That is, we thought that bottom-up grouping processes would often segment simple distractors into whole objects, but we thought that these same processes would often over-segment compound distractors into their constituent parts. Based on this assumption, we reasoned that search stimuli that contain the same number of objects may not contain the same number of search items. In particular, we predicted that because compound distractors comprise more search items than simple distractors, the compound distractors would produce slower search times.

This is indeed what we found. With cluttered displays, search was slower with compound dis-

tractors than with simple distractors. Importantly, when the same stimuli were pre-segmented into a sparse array, no such difference was found. Because this control condition indicated that, once segmented, the simple and compound distractors were equally discriminable from food targets, we attributed the different results for the two distractor types to differences in their segmentability.

If we assume that the critical difference between the simple and compound distractors is their segmentability, then there are at least two accounts of distractor processing that are consistent with the observation that search is slower with compound distractors. In the account mentioned above, distractor items correspond to image regions that are defined by bottom-up grouping processes. Since we assume that compound distractors are segmented into more regions than are simple distractors, we expect compound distractors to produce slower search times. In an alternative account, the distractor items are whole objects. Since whole compound objects cannot be segmented by bottom-up grouping processes, the segmentation of these distractors requires top-down processes. Top-down segmentation is expected to be slower than bottom-up segmentation because it involves iterations between segmentation and recognition. Because the segmentation of compound distractors is assumed to require slow top-down processes, we expect compound distractors to produce slow search times. Note that while both of these accounts predict slow search times for compound distractors, they differ in whether these distractors are processed as whole, recognizable objects or as multiple, smaller parts.

Here we report an experiment that was designed to distinguish between these alternatives. In this experiment, the observer searched for a category target among either normal or chimerical distractors. The normal distractors were familiar objects, while the chimerical distractors were created by swapping parts between the familiar objects. We reasoned that if the items for search are parts defined by bottom-up grouping processes, then search times will be similar for normal and chimerical distractors. On the other hand, if the items for search are whole objects segmented by top-down processes, then search times should differ for the two types of distractors. Recognizing that observers may switch between these two strategies depending on their search task, we conducted this experiment under several conditions.

2 Experiment 1

The purpose of this experiment was to examine whether observers process whole distractor objects when searching through clutter. In choosing the conditions for this study, we adopted the rather simplistic notion that the longer observers look at the display, the more likely they will process the distractors to a deep level, that is, to the level of whole object recognition. With this simple idea in mind, we conducted three versions of the experiment that were expected to produce different search times. In all three versions, the distractors were man-made objects presented in their normal or chimerical form. The versions differed in their targets. In the first version, the search targets were animals in their upright orientation. Because the low-level features of animals may differ from those of the man-made distractors (Levin, Takarae, Miner, & Keil, 2001), we assumed that this target category would produce relatively efficient search. In the second version, the search targets were vehicles, also in their upright orientation. We expected these man-made targets to produce less efficient search. In the third version, we used animal and vehicle targets that were both randomly intermixed and randomly rotated. By intermixing the target categories and by randomly rotating the targets, we increased target uncertainty. The combination of these two additional sources of uncertainty was expected to produce very inefficient search.

2.1 Methods

2.1.1 Search Items

The search items were selected from the Hemera Photo-Object collection and were manipulated using Adobe Photoshop and MatLab. The set of target items consisted of 178 photographs of animals, Figure 1, and 178 photographs of vehicles, Figure 2. In selecting the targets, we tried to sample broadly from the categories. Thus the animals ranged from bees to elephants, and the vehicles ranged from tricycles to tanks.

The set of distractor items consisted of 90 photographs of mostly man-made objects (tools, furniture, office supplies etc.). In selecting the distractors, we applied two criteria in sequence. First, we selected objects that had a salient internal boundary between two regions with different colors or textures. Then, from these objects, we selected pairs that could be combined to create chimeras. Objects were paired only if they had a very similar internal boundary. To exactly match these

boundaries, it was often necessary to rotate or rescale one of the object images. In addition, objects were paired only if they had a similar spatial relationship between their parts. Thus, if the parts were collinear in the donor objects, they remained collinear in the chimerical objects. By preserving the spatial relationship between the parts, we attempted to match the structural grouping cues in the chimerical and normal objects. Once the matches were made, the objects were cut along their boundaries and the parts from paired objects were interchanged, Figures 3 and 4.

2.1.2 Search Displays

The search displays were created in MatLab. To keep the density of the clutter constant across conditions, we scaled the area of the display with the number of search items. The displays were squares of width 210, 297, or 364 pixels, and they contained 8, 16 or 24 search items, respectively.

First, the distractors were added to each display. The distractors were chosen randomly but without replacement from either the normal or chimerical distractor sets, Figure 5 (top row). Before a distractor was added to the display, it was rescaled to have randomly chosen area within the range of 8,000 pixels to 16,000 pixels. The purpose of this random rescaling was to prevent observers from attending to image chunks of a particular size. The distractors were presented in their upright orientation.

Next, a target was selected from one of the categories. Unlike the distractors, the target selection was systematic: each target appeared no more than once in each condition. Before the target was added to the display, it was rescaled in the same way as the distractors. In the single category conditions (only animals or only vehicles), the orientation of the target was upright. In the mixed category condition (animals and vehicles, intermixed), the target was randomly rotated by 0, 90, 180 or 270 degrees. The target's position was also chosen randomly subject to one constraint. Because the observer's task was to identify the display quadrant that contained the target, the target's center of mass was displaced from the vertical and horizontal midlines by at least 20% of the display width. Note that because the target was added last to the displays, it was always completely visible.

For each of the six conditions (3 levels of distractor number \times 2 types of distractor) we generated 178 stimuli (i.e., one stimulus per target). For each observer, we randomly selected 120 stimuli per condition. The 720 stimuli that were selected for each observer were organized into 16 blocks of



Figure 1: Examples of animal targets.



Figure 2: Examples of vehicle targets.



Figure 3: Examples of normal distractors.



Figure 4: Examples of chimerical distractors.

quadrant that contained the target, the observer pressed one of four keys arranged in a square. (The relative position of each key matched its corresponding quadrant.) After a response was registered, there was a one second delay before the next stimulus was presented. Auditory feedback was provided after incorrect responses.

2.1.4 Observers

Eight observers participated in each version of the experiment. None of the observers participated more than once, and none were aware of the purpose of the study. The observers were recruited from the introductory psychology subject pool at Rutgers-Camden. All observers reported having normal color vision and normal or corrected-to-normal acuity.

2.2 Results

In this experiment, observers searched for a category target located among either normal distractors or chimerical distractors. These two distractor types had identical component parts and similar part relationships; the critical difference between them was whether their parts were arranged to create a familiar object. Because we thought that the effect of distractor type might depend on the difficulty of the search task, we conducted three versions of the experiment that were expected to produce different search efficiencies. Despite the range of search times observed in this experiment (Figure 6), there was no evidence of an effect of distractor type under any condition. This result suggests that when observers search for a category target in clutter, they do not recognize whole distractor objects in order to reject them.³

In this first experiment, we varied the difficulty of the search task by varying the target category. We also varied search difficulty by intermixing the target categories and by randomly rotating the targets. None of these search tasks produced a difference between the normal and

³We should mention an exception to this claim. When we first tried this experiment, we included a few dolls in our distractor set. The chimeras created from the dolls included, for example, a vase with a baby doll's head and, as the complementary chimera, flowers coming from a baby doll's body. When we used these doll chimeras in the experiment, they slowed search in every condition. In addition, observers recalled seeing the doll chimeras even though they could not recall any of the other chimeras. We thought it possible that these doll chimeras were slowing search, not because they were unfamiliar, but because the body parts attracted attention and the chimeras were creepy. Whatever the cause, the doll chimeras appear to be an exceptional case, and so we did not use them in the study.

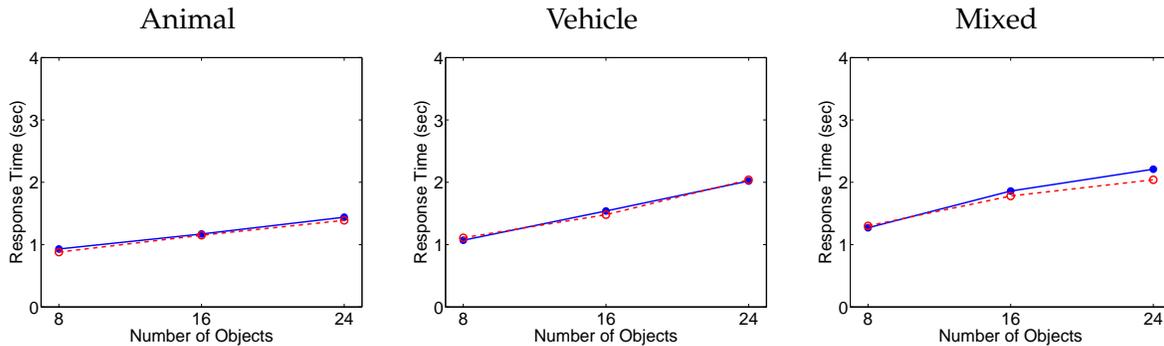


Figure 6: Search times for unoccluded targets (Experiment 1). Left: upright animal targets, middle: upright vehicle targets, right: randomly rotated, animal and vehicle targets. Solid lines: normal distractors, dashed lines: chimerical distractors.

chimerical distractors. Before concluding that even with difficult search tasks, observers do not process distractors as whole objects, we tried another approach to increasing task difficulty. In the next experiment, we examined whether partially occluding the target would produce different search times for the normal and chimerical distractors.

3 Experiment 2

In the previous experiment, the target was added last to the displays and so it was always completely visible. In this second experiment, we added the targets midway during stimulus generation so that it was often only partially visible, Figure 5 (bottom row). Although the placement of the distractors was random, we expected that on at least some of the trials, the distractors would occlude enough of the target to make recognition difficult. If target recognition were sufficiently impaired, we thought observers might find it more expedient to search for the target using a process of elimination. That is, we thought they might determine which regions of the display could be assigned to a distractor and then scrutinize any unassigned regions as a possible target.

3.1 Methods

The search items and the procedure were the same as in the previous experiment. The only difference between the experiments was in the way the stimuli were generated. In this experiment,

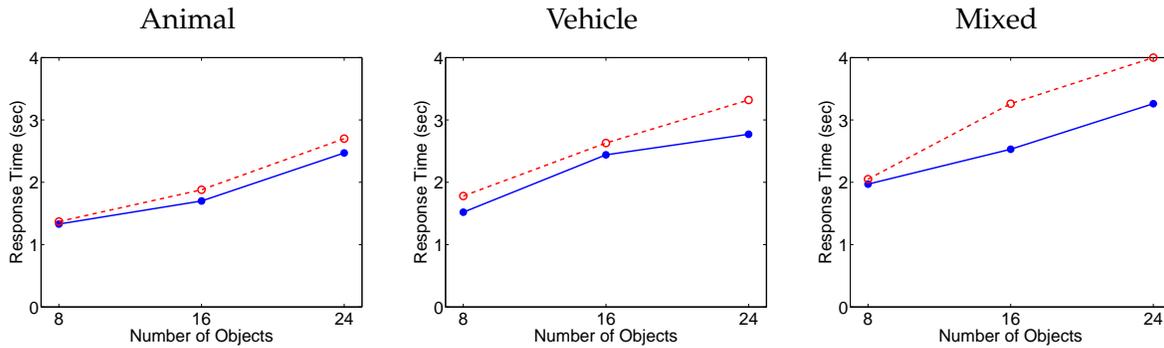


Figure 7: Search times for partially occluded targets (Experiment 2). Left: upright animal targets, middle: upright vehicle targets, right: randomly rotated, animal and vehicle targets. Solid line: normal distractors, dashed line: chimerical distractors.

half of the distractors were added to the display, then the target was positioned in one quadrant, and then the remaining half of the distractors were added. Because the distractor locations were selected randomly, they often occluded the target, and the amount of this occlusion varied considerably across displays. Because a completely hidden target would make the task impossible, we regenerated displays that had excessive amounts of occlusion. Thus, in every display at least 50% of the target was visible.

3.2 Results

In this second experiment, the distractors occluded up to half of the target. Not surprisingly, we found that partially hiding the target makes finding it more difficult, as a comparison of Figure 6 and Figure 7 shows. Of greater interest is the finding that with this more difficult task, there was a difference between the search functions for the normal and chimerical distractors ($F[1, 21] = 57.311, p < 0.000001$). There was also a significant interaction between distractor type and target type ($F[2, 21] = 4.32, p = 0.027$); the greatest difference between the normal and chimerical distractors was seen with the randomly rotated, mixed category targets.

This experiment indicates that when the search target is partially occluded, observers do process the distractors as whole objects. In the discussion section we consider why occluding the target might produce an effect of distractor type.

4 Discussion

In an earlier paper (Bravo & Farid, 2004), we claimed that when observers conduct a serial search of a densely cluttered display, the items of this search do not always correspond to whole objects. Here, we further tested this idea by having observers search for a category target among normal and chimerical distractors. As sets, these two types of distractors were composed of the same parts with the same part relationships. The critical difference between the two distractor types was in whether the parts formed a recognizable object. Our experiment showed that when observers searched for unoccluded targets, whether these targets corresponded to upright animals, upright vehicles or a mixture of randomly rotated animals and vehicles, search times were similar for the normal and chimerical distractors. For these search tasks, observers do not appear to process the distractors as whole objects. In contrast, when observers searched for an occluded target, we found faster search times with the normal distractors. For these search tasks, observers appear to adopt a search strategy that exploits the familiarity of the normal distractor objects. We discuss each of these outcomes below.

The results from our first experiment suggest that when searching for unoccluded targets, observers do not select whole distractors in order to reject them as non-targets. This result would have seemed obvious to our observers. After completing the experiment, we asked observers if they had noticed anything peculiar about the distractors. A typical response was, "I didn't notice anything. You told me to look for the animal, so I didn't look at the other stuff." To the observers it seemed obvious that searching for an animal involves recognizing the animal, not recognizing the distractors. To the observers, the distractors were simply stuff.

The idea that in dense clutter the distractors objects are just stuff is one we wish to emphasize because it contrasts with the idea that the distractors are treated as whole objects. We reported previously that for category search in clutter, search times depend not only on the number of objects in the display but also on the number of object parts. We interpreted this result as evidence that when searching through clutter, observers select and reject parts defined by bottom-up grouping processes. In proposing this interpretation, we acknowledged that the results were also consistent with the idea that slow, top-down processes are needed to segment multi-part distractors while fast bottom-up processes are sufficient to segment single-part distractors. That is, we

acknowledged that the results were also consistent with the idea that observers were selecting and rejecting whole objects. The present experiment refutes this second account: if observers segment multi-part distractors through top-down processes, then we would expect different results for the normal and chimerical distractors. The absence of an effect, coupled with our previous results, suggests that observers searching for an unoccluded category target in clutter do not process the distractors as whole objects.

Our finding of similar search times for the familiar, normal distractors and the unfamiliar, chimerical distractors seems to argue against a familiarity effect in search. But a number of studies have clearly demonstrated such an effect, and so it is important to note the methodological differences between the present study and these earlier studies (Richards & Reicher, 1978; Wang, Cavanagh, & Green, 1994; Lubow & Kaplan, 1997; Shen & Reingold, 2001; Malinowski & Hubner, 2001). In the earlier studies, the familiar items were typically letters that were transformed into unfamiliar items by an image rotation or an image reflection. The same target and distractors were used on every trial, and often all of the distractors in the display were identical. Additionally, the letters were sparsely distributed, so there was never any spatial overlap between items. And finally, because observers had likely been exposed to the items millions of times prior to the experiment, the items were exceedingly familiar. Each of these task characteristics: item certainty, distractor homogeneity, lack of clutter, and over-learned items, could be critical for an effect of distractor familiarity.

With our task of category search in clutter, we found an effect of familiarity only when the target was partially occluded. While additional experiments are needed to delineate the conditions that produce the familiarity effect, we propose that any factor that severely degrades target recognition may cause observers to adopt a search strategy that involves recognizing whole distractor objects. That is, we propose that when an initial scan of the display fails to turn up a target, observers may re-examine the display using a search strategy that involves assigning each region in the display to a distractor. If some regions cannot be so assigned, then they may be evaluated as potential targets (Bravo & Farid, 2003).

This process of elimination will be most successful with familiar distractors for two reasons. First, although some individual distractor parts may be recognizable, other distractor parts may not be. When these indistinctive parts belong to familiar objects, observers can use object knowl-

edge to rule them out as possible targets. And second, the process of mentally eliminating distractors involves working memory, which is known to have a limited capacity (Luck & Vogel, 1997; Alvarez & Cavanagh, 2004). Observers may be able to hold more distractor parts in working memory when the parts are organized as familiar objects.

An observer who casually inspects our normal, cluttered displays perceives a collection of recognizable objects. Although recognition of the objects occurs automatically and effortlessly, it reflects a process that vision researchers find so perplexing that it currently seems inexplicable. The reasons why object recognition is so problematic are well-known and include such factors as viewpoint, lighting and exemplar variability. In our displays, these difficulties were further compounded by inconsistencies in scale and by extreme clutter. Given the difficult nature of object recognition, it is unsurprising that, whenever possible, observers adopt search strategies that minimize the involvement of recognition. We may perceive a world of discrete objects, but when we are engaged in search, it is all just stuff.

Acknowledgments

We thank J. William Whitlow for performing the statistical analysis of our data.

5 References

- Alvarez, G., & Cavanagh, P. (2004). The capacity of visual short term memory is set both by visual information load and by number of objects. *Psychological Science, 15*, 106–111.
- Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.
- Bravo, M., & Farid, H. (2003). Object segmentation by top-down processes. *Visual Cognition, 10*, 471–491.
- Bravo, M., & Farid, H. (2004). Search for a category target in clutter. *Perception, 33*, 643–652.
- Bravo, M., & Farid, H. (2006). Object recognition in dense clutter. *Perception and Psychophysics, in press*.
- Levin, D. T., Takarae, Y., Miner, A., & Keil, F. C. (2001). Efficient visual search by category: Speci-

- fying the features that mark the difference between artifacts and animals in preattentive vision. *Perception and Psychophysics*, 63, 676–697.
- Lubow, R., & Kaplan, O. (1997). Visual search as a function of type of prior experience with targets and distractors. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 14–24.
- Luck, S., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Malinowski, P., & Hubner, R. (2001). The effect of familiarity on visual-search performance: evidence for learned basic features. *Perception and Psychophysics*, 63, 458–463.
- Pelli, D. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Richards, J., & Reicher, G. (1978). The effect of background familiarity in visual search. *Perception and Psychophysics*, 23, 499–505.
- Shen, J., & Reingold, E. (2001). Visual search asymmetry: The influence of stimulus familiarity and low-level features. *Perception and Psychophysics*, 63, 464–475.
- Ullman, S. (1997). *High-level vision: Object recognition and visual cognition*. Bradford/MIT Press.
- Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception and Psychophysics*, 56, 495–500.